



Department for
Communities and
Local Government



English Housing Survey

Quality Report 2014-15



Contents

Summary

Introduction

- [Primary purpose](#)
- [A brief history of the English Housing Survey](#)
- [Overview of methodology](#)
- [English Housing Survey quality review](#)

Relevance

- [Annual user consultation](#)
- [User group membership](#)
- [Reporting strategy for 2014-15](#)
- [Publication consultation on the future shape of the EHS](#)

Accuracy and reliability

- [Sampling error](#)
- [Coverage error](#)
- [Measurement error](#)
- [Processing error](#)
- [Response rates](#)
- [Model assumption error](#)

Timeliness and punctuality

Accessibility and clarity

- [GOV.UK](#)
- [UK Data Service End User Licence](#)
- [UK Data Service Special Licence](#)
- [UK Data Service Secure Access](#)
- [Open data standards](#)
- [English Housing Survey data security strategy](#)

Coherence and comparability

- [Devolved administration data sources](#)
- [Geographies below England level](#)
- [Harmonised questions](#)

Sources for further information or advice

Summary

This Quality Report provides a brief summary of key issues relating to quality that users of the English Housing Survey (EHS) outputs need to be aware of. The report begins with purpose and background of the EHS, a brief introduction to methodology, and a summary of work undertaken in 2015 and 2016 to improve the quality of EHS statistics. A more detailed discussion of survey methodology and how it impacts on quality is provided in the EHS Technical Report, which includes chapters on sampling, questionnaire, fieldwork, response rates, data processing, weighting and standard errors. Note that the format of this Quality Report differs from the Technical Advice Note on Data Quality published up to 2011-12, which focused solely on the accuracy quality dimension.

This report covers all dimensions of quality defined by the European Statistical System as recommended by the Code of Practice for Official Statistics.

- **Relevance** – This section discusses work undertaken to ensure that the EHS meets user needs, including details of the user consultation process for the 2014-15 reporting strategy and the 2016-17 questionnaire.
- **Accuracy and reliability** – This section summarises the main sources of bias and other errors in a sample survey of this type, which impact on the degree of closeness between EHS estimates and true population value.
- **Timeliness and punctuality** – This section provides information on headline and all other data releases from the EHS, and considers trade-off between timeliness and the other quality dimensions.
- **Accessibility and clarity** – This section covers different ways of accessing EHS outputs, including reports and tables available on Open Government Licence on gov.uk, End User Licence, Special Licence and Secure Access datasets at the UK Data Service.
- **Coherence and comparability** – This section explains how the EHS monitors and reports on coherence and comparability of survey over time, harmonisation of the survey with ONS standards, reporting of geographic variables, and information on similar housing surveys carried out by devolved administrations.

1. Introduction

Primary purpose

- 1.1 The EHS is the Department for Communities and Local Government (DCLG) flagship survey of people's housing circumstances and the energy efficiency and condition of the housing stock in England. It covers all housing tenures and provides valuable information and evidence to inform the development and monitoring of DCLG's housing policies.
- 1.2 The Department for Business, Energy and Industrial Strategy (BEIS) also make an annual financial contribution to the running cost of the survey and use the data to measure the effectiveness of its policies designed to improve the energy efficiency of English homes and to monitor fuel poverty. The Department for Business, Energy and Industrial Strategy was created on 14 July 2016. This Department took over energy policy from the former Department of Energy and Climate Change. Within this publication references to BEIS's predecessor Department refer to DECC.

A brief history of the English Housing Survey

- 1.3 The EHS was launched in April 2008 bringing together two former housing surveys – the Survey of English Housing and the English House Condition Survey.
- 1.4 The English House Condition Survey started in 1976 and was carried out every five years until 2001. From 2003 onwards, the English House Condition Survey operated continuously until the merger to form the EHS. Prior to the English House Condition Survey, a Regional Housing Survey in 1967 and Housing Condition Survey in 1971 included a physical survey similar to that in the English House Conditions Survey. The English House Condition Survey reports are available on The National Archives website.¹
- 1.5 The Survey of English Housing was a continuous household survey that collected information from nearly 20,000 households each year about the characteristics of their housing and their attitudes to housing and related issues. It operated continuously from 1993-94 and the final fieldwork year for the Survey

¹<http://webarchive.nationalarchives.gov.uk/20121108165934/http://www.communities.gov.uk/housing/housingresearch/housingsurveys/englishhousecondition/>

of English Housing was 2007-08. The Survey of English Housing reports are available on The National Archives website.²

- 1.6 The first four English Housing Surveys (2008-09 to 2011-12) were conducted by the Office for National Statistics, MMBL-CADS and the Building Research Establishment (BRE). A consortium consisting of NatCen Social Research, CADS Housing Surveys, BRE and Bryson Purdon Social Research carried out the 2012-13 to 2015-16 surveys. In 2015, DCLG awarded a new contract covering the survey years 2016-17 to 2018-19 to NatCen Social Research in consortium with CADS Housing Surveys and BRE.

Overview of methodology

- 1.7 The population or key units of interest in the EHS are residential households and dwellings in England. A dwelling is defined as a unit of accommodation where all the rooms and amenities are for the exclusive use of the household(s) occupying them. The EHS collects information about the respondents' main home so it defines a household as one person or a group of people (not necessarily related) who have the accommodation as their only or main residence, and (for a group) share cooking facilities and share a living room or sitting room or dining area. This definition is slightly different from the definition used in the 2011 Census.³
- 1.8 The EHS has a complex two-stage methodology, including a household interview in occupied dwellings and a physical inspection of occupied and vacant dwellings by a qualified surveyor. The survey is continuous, operating in four quarters (eight waves) over the year. However, to accommodate a delay to the start of the 2015-16 survey caused by the consultation on the future of the EHS, fieldwork for that year took place over three quarters of the year, beginning in July 2015.
- 1.9 A household interview is undertaken with approximately 13,300 households each year (referred to as 'full interview sample'). These data are analysed on an annual financial year basis, with the fieldwork for the 2014-15 survey year completed between April 2014 and March 2015.
- 1.10 Around 200 of the dwellings identified as vacant in the interview survey and the dwellings of 6,000 of the households taking part in the interview survey were randomly selected to take part in the physical inspection. This group is referred to as the 'dwelling or household sub-sample' depending on whether the analysis

²<http://webarchive.nationalarchives.gov.uk/20121108165934/http://www.communities.gov.uk/housing/housingresearch/housingsurveys/surveyofenglishhousing/>

³ The definition of a household in the 2011 Census was one person living alone, or a group of people (not necessarily related) **living at the same address** who share cooking facilities and share a living room or sitting room or dining area.

relates to dwellings or households. These data are analysed on a two year rolling basis with the results presented for '2014' based on fieldwork conducted between April 2013 and March 2015.

- 1.11 The sample of the EHS is representative of the population of England. For the survey years 2008-09 to 2011-12, the sample was selected using a simple random sample design. The sample design changed in 2012-13 so that the sample is currently unclustered over two years of the survey, but is clustered over a single year. The advantage of this approach is that it reduces the fieldwork area to half the country so interviewer and surveyor travel time/costs are significantly reduced.
- 1.12 The EHS sample is stratified by geographic region, tenure and households with an HRP who worked in non-manual occupations (2001 Census), and subsampled by tenure (identified using Experian's Residata⁴).
- 1.13 To produce unbiased estimates for the population of households and dwellings in England, a process of weighting is carried out. This includes:
 - scaling up to the total number of addresses in sample frame;
 - adjusting for selection probabilities where there is not a one-to-one relationship between the address and dwelling or the dwelling and household;
 - models to produce response weights adjusting for non-response bias;
 - calculating a design factor to adjust for tenure subsampling; and
 - calibration to control totals.
- 1.14 From 2013-14, the weights were calibrated to one set of control totals for households/dwellings using the DCLG dwelling counts. Before that, the calibration was to two sets of control totals for households by age, sex and tenure from the ONS Labour Force Survey and dwelling control totals from DCLG dwelling stock. Details on the weighting methodologies applied before and after 2013-14 are available from: <https://www.gov.uk/government/collections/english-housing-survey-technical-advice>
- 1.15 The EHS methodology is detailed further in the technical report published in July 2016.⁵

⁴ Experian possess a database that contains information obtained from a number of sources including insurance companies, Census, etc. referred to as Residata. It is from this that information is taken on predominant tenure within a postcode as well as other information. The matching of the EHS sample to Residata is carried out by BRE.

⁵ <https://www.gov.uk/government/collections/english-housing-survey-technical-advice>

English Housing Survey quality review

- 1.16 The EHS team have undertaken several actions over the course of the year to improve accuracy and accessibility of the EHS data. This included an internal error log put in place to record revisions made and how errors arose, and monitoring to improve processes to reduce errors arising. A series of training events was held in 2016 to enhance capability within DCLG for analysing and quality assuring analysis of the survey. An explanatory note is being produced for the department on the quality of income data collection and which of the EHS income measures to use for different purposes.
- 1.17 The EHS team participated in a cross government surveys group: to share knowledge, experiences and good practice with others; and to offer a chance to discuss and raise awareness of current issues. The Style Guide for the EHS was updated to reflect new guidance from the GSS Good Practice Team on report writing and data visualisation (see Section 2 User Consultation for further details on changes to dissemination strategy).
- 1.18 The published technical documentation was updated to enhance the accessibility of the survey datasets. This included production of data dictionaries of the detailed derived datasets⁶ and further annexes in Chapter 5 of the 2013-14 technical report, covering the production of household type and dimensions derived variables (see Section 5 for more information on release of datasets on UK Data Service).
- 1.19 There was also a review of the process for taking account of sampling error in the presentation of the EHS data, led by the Technical Advisory Group. After considering two options for conducting significance testing in the headline and annual reports, the group agreed to continue the existing approach, as indicated in the first bullet:
- conduct significance testing using formula below with an average design factor of 1.2 for exploratory analysis and consider calculating specific values using specialist software if borderline significance;

$$\text{difference} = p1 - p2$$

$$\text{Standard error of difference (simple random sample)} = \sqrt{\frac{p1 \times (1 - p1)}{n1} + \frac{p2 \times (1 - p2)}{n2}}$$

⁶ http://doc.ukdataservice.ac.uk/doc/7802/mrdoc/pdf/7802_detailed_derived_variable_data_dictionaries.pdf

$$\text{calculated value} = \left(\frac{\text{difference}}{\text{standard error of difference}} \right) \div \text{design factor}$$

if the calculated value > or = 1.96 the difference is significant at <0.05 95% level

Note: p = proportion n = sample size

- all significance testing using STATA or SPSS complex samples commands/modules taking account of complex sample design in calculation of standard errors and design factors.
- 1.20 To provide users with an indication of the sampling error calculated using the second option, DCLG publish standard errors, design factors and confidence intervals for several key variables in Chapter 7 of the technical report. These outputs are currently computed in STATA, with an additional adjustment in Excel for Finite Population Correction at the area level (half of all clusters are selected each year, which results in the intra-cluster correlation being halved in size).
- 1.21 EHS data validation has been consolidated and further automated by bringing the syntax necessary for the various data checking and editing stages into one Excel workbook, from which the validation process can easily be managed and maintained.

2. Relevance

The degree to which statistical outputs meet users' needs⁷

Annual user consultation

- 2.1. Each year, the EHS team conducts a review of forthcoming data collection and reporting priorities to ensure that user requirements are being met. The user consultation carried out in Autumn 2015 had three main objectives:
 - to review the data collected through the interview survey questionnaire to ensure that policy requirements were being met;
 - to review the modelled data to ensure that the concepts measured and standards applied are still appropriate to what users want to measure; and
 - to review the outputs published, in particular the annual reports and the live tables.
- 2.2. The EHS team organised a number of meetings with policy teams and analysts across DCLG and BEIS.
- 2.3. On the basis of these discussions, the EHS team devised the 2014-15 EHS reporting strategy and a document outlining proposed changes to the 2015-16 questionnaire, both of which were presented to the EHS user group meeting for sign off. No changes were required to the modelled data.
- 2.4. The reporting strategy proposed significant changes to the annual reports, which were signed off by the user group. More details are below.
- 2.5. Few changes were made to the questionnaire and the content of the physical survey was not reviewed. This was due to the late running of the 2016-17 onwards procurement exercise, which meant that there wasn't time to implement changes to the physical survey or make anything more than minor changes to the interview survey in time for the commencement of 2016-17 fieldwork.

User group membership

- 2.6. Membership of the EHS user group is reviewed and refreshed annually to ensure that all relevant DCLG and DECC policy teams have representation. The group is presently comprised of 25 policy officials and 10 analysts. The group is chaired

⁷<http://www.ons.gov.uk/ons/guide-method/best-practice/gss-best-practice/gss-quality-good-practice/quality-measurement-and-reporting-guidance/index.html>

by a principal analyst within the EHS team. The survey's SRO is consulted before any major change is implemented.

Reporting strategy for 2014-15

2.7. The reporting strategy for 2014-15 represents a significant departure from previous years. Rather than produce four lengthy annual reports (on households, profile of English housing, fire and fire safety and energy), a series of nine short topic-based reports was produced on:

- Smoke alarms in English homes
- Housing stock
- The private rented sector
- The social rented sector
- First time buyers and potential home owners
- Adaptations and accessibility
- Energy
- Housing and well-being
- Housing for older people

2.8. This revised reporting strategy was signed off by the user group and was a direct response to feedback from survey users that the existing EHS reports can be too lengthy and technical. This approach also fits well with guidance produced by the GSS good practice team and it is our view that it will make the EHS data more accessible to a wider audience. It was anticipated that each report would contain around 20 pages of content.

2.9. These annual reports supplement information provided in the headline report which was produced by the EHS team and which was published in February 2016. While the user consultation helped to identify priorities for the headline report, the EHS team did not consult directly on its content.

Public consultation on the future shape of the EHS

2.10. DCLG carried out a public consultation on the future shape of the EHS in 2015⁸. The consultation revealed the EHS to be a highly valued source of evidence for policy development at national and local level. In its response to the public consultation, DCLG announced that the most cost effective way forward for the EHS was to continue to carry out the survey on an annual basis.

⁸ <https://www.gov.uk/government/consultations/consultation-on-the-future-shape-of-the-english-housing-survey>

3. Accuracy and reliability

The degree of closeness between an estimate and the true value⁷

- 3.1 This section summarises the steps taken to quality assure the EHS data and the main sources of bias and other errors that impact on a sample survey of this type. The main sources of error are sampling error and non-sampling error, which includes coverage error, non-response error, measurement error, processing error and model assumption error. While it is not possible to measure all sources of error, quantitative estimates are produced where possible and published in the technical report for the 2014-15 survey.

Sampling error

Sampling error is the error that arises because the estimate is based on a sample survey rather than a full census of the population⁷

- 3.2 The results obtained for any single sample may, by chance, differ from the true values for the population but the difference would be expected to average to zero over a number of repeats of the survey. The amount of variation depends on the size of the sample and the sample design and weighting method.
- 3.3 Although the estimates produced from a sample survey will rarely be identical to the population value, statistical theory allows us to measure the accuracy of any survey result. Standard errors can be estimated from the results obtained for the sample, and these allow calculation of confidence intervals which give an indication of the range in which the true population value is likely to fall.
- 3.4 A measure of the impact of the variation introduced by the sample design and the weighting is the design factor (deft). This is evaluated relative to the error that would have been produced had the survey been carried out using a simple random sample. A deft greater than one shows that the design and weighting have increased the variability of the estimate and increased the measure of the standard error relative to simple random sample design.
- 3.5 Sampling error is discussed in more detail in Chapter 7 of the technical report. This includes examples of standard errors and confidence intervals calculated using the appropriate design factors.

Coverage error

Coverage error arises from failure to cover adequately all members of the population being studied⁷

-
- 3.6 The sample is drawn from the Postcode Address File (PAF), a list of addresses maintained by the Post Office. This source provides good coverage of the population of interest, i.e. residential households and dwellings in England. The achieved sample of dwellings does not include any new dwellings built since the creation dates of PAF files from which the sample taken. The weights are adjusted using the number of new dwellings built between the PAF date and the reference date for the weighting to account for this error in coverage.

Measurement error

Measurement error may arise due to inaccuracies in individual measurements of survey variables because of the inherent difficulties of observing, identifying and recording what has been observed⁷

- 3.7 Measurement error may occur randomly, or may reflect a problem experienced by most or all interviewers or surveyors. The key mechanisms in place to minimise measurement error focus on the questionnaire and physical survey form development process and interview and surveyor training.
- 3.8 For the interview survey, NatCen use in-house expertise in question design and testing to ensure that the questions are understood by respondents in the way intended. Full-scale piloting of questions is not undertaken for the EHS. New questions introduced to the survey are reviewed by NatCen and DCLG after the first quarter of data collection.
- 3.9 For the physical survey, BRE and CADS Housing Surveys work together to apply expert knowledge of buildings research and housing surveys to minimise inconsistencies in the data collection process. There are more practical difficulties in assessing the condition of an individual dwelling than the characteristics of a household. These difficulties mainly stem from the technical problems in the diagnosis and prognosis of any defects found in the dwelling. Difficulties are found particularly in more subjective assessments such as the state of repair.
- 3.10 To assess effects of surveyors making different judgements about the same information, a calibration workbook exercise and a surveyor variability study (SVS) were conducted in 2009-10. The calibration workbook was a desk-based exercise, with surveyors asked to record assessments as they would in the field for a set of examples with descriptions and photographs of a number of dwelling faults. Results from the 2009-10 exercise showed no significant difference overall in the surveyors' assessments. The SVS involved a call-back exercise in which 300 properties were re-surveyed by a second surveyor and results compared with the first surveyor. The survey measures with low levels of agreement tended to be found in the topic areas covering external environments and stock

condition, and typically required a surveyor's opinion on topics with generally high degrees of variability (e.g. an opinion of the local area or the condition of a property). The annual training sessions for surveyors have been updated to target more the topics identified as having a higher observed level of variability. The calibration workbook exercise and SVS was repeated in the 2014-15 survey year. Analysis of the data is still ongoing.

- 3.11 All new interviewers receive distance learning material and complete a one-day training course in administering the EHS interview before starting fieldwork. All new surveyors receive distance learning material and complete a residential briefing course before starting work on the survey. Ongoing refresher distance learning via a newsletter and briefing days for interviewers and surveyors are provided as needed to ensure that fieldwork staff are kept up to date with new developments in the EHS.

Processing error

Processing error includes errors in data capture, coding, editing and tabulation of the data as well as in the assignment of weights⁷

- 3.12 Processing error can be systematic, for example, an error in the programming syntax that leads to a wrong code being consistently applied in a particular set of circumstances. Mechanisms in place to ensure accurate data capture and processing and minimise processing error include:
- NatCen carry out systematic testing of the CAPI programme used by interviewers to record interview survey answers
 - BRE conduct extensive testing of the validation systems used in processing physical survey data
 - NatCen and BRE test syntax for automated data checking processes
- 3.13 Processing error can also impact on variance, for example, random keying errors in entering data which, across replications, would cancel each other out. The key checks in place in NatCen and BRE to minimise this source of error include:
- minimising data entry errors by interviewer or incorrect answers given by respondents with a number of validation checks built in to the CAPI programme, for example, extreme numeric values or unlikely combinations of answers
 - review of multi-household or dwelling addresses to ensure selection protocol followed

-
- comparison of key variables collected at household interview with physical survey
 - data outputs are validated with checks such as:
 - case completeness
 - variable completeness
 - investigation of outliers
 - timeseries comparison
 - comparison with external data sources
 - selective case by case analysis

3.14 The number of errors and required alterations to raw data is monitored to ensure that quality standard is maintained.

Response rates

Non-response bias may occur if non-respondents differ from respondents. Non-response can lead to an increase in the variance of survey estimates, as non-response will reduce the sample size⁷

- 3.15 The effect of non-response bias is minimised by steps applied in the weighting process. Valid but non-responding cases are checked to assess if they are typical of those that remain and, if not, to counter any resulting response bias in the grossed data set. Where non-response biases were found at any stage of the survey, adjustments were made to the responding cases in the weighting procedures for that stage.
- 3.16 To ensure that an adequate sample size is attained with an acceptable level of variance, a target number of interview surveys are set and an assumption made for non-response rate, from which the size of the required issued sample is calculated. Response rates are continually monitored and interviewers are trained to maximise response rates.
- 3.17 For item non-response, imputation is carried out when creating key derived variables, using either external data sources or sample median from other information collected in EHS. Addressing non-response through imputation can lead to the appearance of the variance being reduced, as imputed values are usually less extreme than would be observed from sampling alone.
- 3.18 Imputation rates monitored and reported on through the Quality Monitoring reports as part of regular progress meetings with consortium. Overall response rates, key item non-response and imputation rates are published in the EHS technical report.

Model assumption error

- 3.19 In October 2012, the Cabinet Secretary and the Head of the Civil Service commissioned a review of the quality assurance of analytical models that inform government policy. The review published a final report in March 2013, setting out results of work to map business critical models and quality assurance in government.⁹
- 3.20 The models used in the EHS have been reviewed in light of the recommendations in this report and updated model documentation has been produced. In particular, the underlying assumptions of the model are now clearly identified, the inputs required and the key outputs of the models are documented, and any changes to the model are recorded. The model documentation is consistent with that used for other models in DCLG developed as part of a Quality Assurance working group. The model assumptions are discussed further in the EHS technical report.

⁹https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/206946/review_of_qa_of_govt_analytical_models_final_report_040313.pdf

4. Timeliness and punctuality

Timeliness refers to the lapse of time between publication and the period to which the data refer. Punctuality refers to the gap between planned and actual publication dates. Whether the outputs are up to date with respect to users' needs⁷

- 4.1 There is a trade-off between timeliness and the other quality dimensions, in particular accuracy, accessibility and clarity. It is important to ensure that there are adequate processes to ensure accuracy of the large survey dataset and produce clear publication tables, and apply appropriate disclosure-control to public datasets released.
- 4.2 To provide timely data to users, key headline figures from the EHS are published in a headline report, usually around 11 months after the end of the reference period. For the 2014-15 survey year, the headline report was published on 18th February 2016. The publication date for headline reports and further annual reports is pre-announced on the official statistics release calendar <https://www.gov.uk/government/statistics/announcements>
- 4.3 A further series of reports were published on 21 July 2016: Smoke alarms in English homes, Housing stock, The private rented sector, The social rented sector, First time buyers and potential home owners, Adaptations and accessibility, Energy, Housing and well-being, Housing for older people. The 2014-15 End User Licence datasets were released on the UK Data Service in July 2016. The Special Licence and Secure Access datasets will be released on the UK Data Service in 2016 after disclosure control is applied.
- 4.4 BEIS published a report on fuel poverty based on the 2014-15 survey year in June 2016 after the release of the headline report, as specified in the [Memorandum of Understanding in Relation to Data Sharing](#) between DCLG and BEIS (the document refers to DECC as it was set up prior to machinery of government change; this will be updated in due course).
- 4.5 In accordance with Pre-release Access to Official Statistics Order 2008, ministers and eligible staff are given pre-release access to EHS statistics 24 hours before release. The EHS pre-release access list is published on the DCLG statistics web page.¹⁰
- 4.6 The EHS data production and publication schedule are kept under review and will take into account user needs when considering the timeliness of future data releases.

¹⁰ <https://www.gov.uk/government/organisations/department-for-communities-and-local-government/about/statistics>

5. Accessibility and clarity

Accessibility is the ease with which users are able to access the data, also reflecting the format in which the data are available and the availability of supporting information.

Clarity refers to the quality and sufficiency of the release details, illustrations and accompanying advice⁷

5.1 The different formats available for accessing EHS data are described below.

GOV.UK

5.2 The EHS webpages are accessible from the DCLG statistics launch page¹¹. The table below shows the total number of hits and unique users on selected EHS webpages from 1 April 2015 to 31 March 2016.

Number of hits on www.gov.uk between 1 April 2015 and 31 March 2016

Page	Pageviews	Unique Pageviews
government/collections/english-housing-survey	39,187	26,611
/government/statistics/english-housing-survey-2013-to-2014-headline-report	13372	10417
/government/statistics/english-housing-survey-2013-to-2014-household-report	12186	8037
/government/statistics/english-housing-survey-2014-to-2015-headline-report	4876	3678
/government/statistics/english-housing-survey-2013-profile-of-english-housing-report	3213	2182
/government/statistics/english-housing-survey-2013-energy-efficiency-of-english-housing-report	1579	1074
/government/statistics/english-housing-survey-2013-to-2014-fire-and-fire-safety-report	673	460

5.3 Feedback from users indicates that it is important for them to be able to find statistics quickly. Following this feedback, the EHS team reviewed and updated the EHS webpages in July 2016. To improve the accessibility of statistics links to all findings and information on the UK Data Service datasets can now be directly accessed from the frontpage. In addition, the questionnaire and survey form for

¹¹ <https://www.gov.uk/government/organisations/department-for-communities-and-local-government/about/statistics>

each year can be found with the reports from that year, to make them more easily accessible. A new dedicated page for households selected to participate in the survey has also been set up, with a direct link from the frontpage. Over the next year the EHS team will continue reviewing the webpages. The address of the frontpage is <https://www.gov.uk/government/collections/english-housing-survey>.

- 5.4 Between August 2015 and end of March 2016 the EHS team received more than 130 queries from outside the government, more than 20 queries from other government departments (excluding BEIS) and more than 70 request for ad-hoc analysis by DCLG colleagues. The team also responded to a large number of Parliamentary Questions.

UK Data Service End User Licence

- 5.5 The anonymised EHS datasets are published every year, via End User Licences available to all members of the public at the UK Data Service <http://ukdataservice.ac.uk/>. Beginning with 2014-15 data, simplified versions of the derived datasets will be available as End User Licence datasets, while the more detailed data will be available as Special Licence datasets. The lowest geography in the End User Licence datasets will be region.
- 5.6 The registration process for access to the UK Data Service is different for UK university or college users than for other UK users, non-UK users and commercial users. This is explained in detail here: <http://ukdataservice.ac.uk/get-data/how-to-access/registration.aspx>
- If you are based at a UK institution of higher or further education (UK HE/FE), you will need to contact your library, IT helpdesk or e-resources administrator. Your institutional web pages or intranet may also contain useful information on how to obtain a username and password.
 - For other UK users, non-UK users and Commercial users, it is a two-stage process. The first step is to apply for a UK Data Archive username and password at: <http://www.data-archive.ac.uk/sign-up/credentials-application>
 - Once these have been received, complete the registration process as detailed for UK University or college users, selecting 'UK Data Archive' as organisation when you log in and/or register.
- 5.7 The datasets are available in SPSS, Stata or tab formats. Due to limited demand, data are not routinely made available in SAS or R, but data in SPSS format can be imported in some versions of these software. Any potentially disclosive information is removed and only high level statistical region geographic identifiers

are included. Checks are undertaken to avoid the release of disclosive information via for example string variables.

- 5.8 For any problems accessing EHS datasets on UK Data Service, please contact the UK Data Service directly at <http://ukdataservice.ac.uk/help/get-in-touch.aspx>. UK Data Service will refer any queries that they cannot answer to the EHS team at DCLG.
- 5.9 From 1 April 2015 to 31 March 2016, EHS End User Licence datasets have been downloaded from the UK Data Service 827 times by 211 User IDs for 224 unique projects.¹² The majority of users are staff and students at higher and further education institutes. Other users include charities, not-for-profit organisations, central and local government, and commercial users (for non-commercial purposes only).

UK Data Service Special Licence

- 5.10 Beginning with 2014-15 data, the more detailed datasets will only be available as Special Licence datasets. The Special Licence datasets will include the detailed 'raw' interview and 'raw' physical survey files and extended versions of the derived datasets. For further detail, a list of the general, interview and physical derived variables released under the End User Licence and Special Licence is available in Chapter 5 of the EHS technical report.
- 5.11 This decision was taken in order to increase the disclosure control on the datasets to comply with new Government Statistical Service guidance contained in GSS/GSR Disclosure Control Guidance for Microdata Produced from Social Surveys.¹³ The content and the level of disclosure control applied in the 'new' Special Licence datasets is similar to what had been released under the End User Licence up to 2013-14. For more information on Special Licence datasets, see <https://www.ukdataservice.ac.uk/get-data/how-to-access/conditions/special-licence>.
- 5.12 The initial registration process is the same as for EUL datasets (see para 5.6). Users will then fill in a form, providing further information on their research training and background as well as the intended use of the data. Anyone with a good level of research training and/or practical experience can be granted access to the Special Licence datasets, including users at local authorities, charities or independent researchers. For further information on the application

¹² These figures exclude the Fuel Poverty Datasets. Including those, the figures are as follow: From 1 April 2015 to 31 March 2016, End User Licence datasets have been downloaded from the UK Data Service 937 times by 229 User IDs for 247 unique projects.

¹³ Published October 2014, available from <http://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/disclosurecontrol/policyforsocialsurvey/microdata>

process, see the tab for 'Order Special Licence datasets' on the UK Data Service webpage: <https://www.ukdataservice.ac.uk/get-data/how-to-access/downloadorder/ordersl>.

UK Data Service Secure Access

- 5.13 The UK Data Service Secure Access provides access to data that are too detailed, sensitive or confidential to be made available under the standard End User Licence or Special Licence. Data accessed in this way cannot be downloaded. Once researchers are specially trained, they analyse the data remotely from their institutional desktop or in a Safe Centre at University of Essex.
- 5.14 The initial registration process is the same as for accessing End User Licence datasets, with a more detailed application process once logged on, details here: <https://www.ukdataservice.ac.uk/get-data/how-to-access/accessecurelab>
- 5.15 The EHS datasets with detailed geographic identifiers are available only via Secure Access at the UK Data Service.

Open data standards

- 5.16 The Open Data White Paper '*Unleashing the Potential*' published by Cabinet Office in June 2012 announced that 'the Government intends to adopt the Five Star Scheme as a measure of the usability of its Open Data.'¹⁴
- 5.17 The data published on EHS pages of GOV.UK is subject to rights detailed in the Open Government Licence v2.0, as specified on the DCLG statistics summary page: 'All content is available under the <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/2/> , except where otherwise stated'. The data is published in pdf format and Excel (Level 2 available on web with an open licence in proprietary machine-readable format)
- 5.18 The more disclosive nature of the datasets at the UK Data Service dictates a more restrictive licence than the Open Government Licence. As discussed above, there are datasets released on an End User Licence, Special Licence and a dataset with detailed geographic identifiers released with Secure Access licence. While not meeting level 1 of open licence because of the necessary disclosure control restrictions, the datasets from this service are available in tab-delimited format (Level 3 criteria of non-proprietary, machine-readable).

¹⁴https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/78946/CM8353_acc.pdf

-
- 5.19 The DCLG Linked Open Data Communities is an initiative to improve accessibility of datasets held by the Department.¹⁵ The EHS is not currently included in the sets of data on Open Data Communities, but will be considered as the site is expanded to publish new data. The Open Data Communities is Level 5, providing linked data in multiple machine-readable formats including JSON, RDF, Turtle and N-triples, and a SPARQL 1.1 endpoint.

English Housing Survey data security strategy

- 5.20 In July 2013, a review of the data security strategy for EHS was undertaken and a new agreement published on EHS webpages.¹⁶

¹⁵ <http://opendatacommunities.org/>

¹⁶ <https://www.gov.uk/government/publications/english-housing-survey-data-security-strategy-and-arrangements>

6. Coherence and comparability

The degree to which data can be compared over time and domain (for example geographic level). The degree to which data that refer to the same topic but are derived from different sources or methods, are similar⁷

- 6.1 Following user feedback, the annual reports will expand analysis on long-term trends and information on changes in definitions over time in the English House Conditions Survey, Survey of English Housing and EHS. Variable naming has been retained from English House Conditions Survey and Survey of English Housing variables where there has been no significant change to the data collected.
- 6.2 DCLG routinely compare EHS data outputs with comparable data sources on the same topics. Where inconsistencies are identified, these will be flagged up in our reports.
- 6.3 The EHS team reviewed and changed the weighting methodology in 2013-14. The new weighting approach may have an impact on the coherence with other available data sources on households and dwelling stock and on the comparability of the EHS data over time.

Devolved administration data sources

- 6.4 Scotland, Northern Ireland and Wales conduct their own housing surveys. There has been limited work in recent years by DCLG on evaluating if outputs have been produced in comparable ways as for the EHS. Users are advised to review technical advice provided for each of the surveys to assess how variables of interest in their analysis compare across each of the four surveys.
 - **Scottish House Condition Survey** – the largest single housing research project in Scotland, and the only national survey to look at the physical condition of Scotland's homes as well as the experiences of householders. The survey is now an integrated component of the Scottish Household Survey since 2012.¹⁷
 - **Northern Ireland House Condition Survey** – provides a comprehensive picture of the dwelling stock, including condition and energy efficiency, and

¹⁷ <http://www.scotland.gov.uk/Topics/Statistics/SHCS>

examine the association between dwelling conditions and the social and economic circumstances of households.¹⁸

- **The Living in Wales survey** – carried out from 2004 to 2008, and featured a property survey by a qualified surveyor in 2004 and 2008.¹⁹ This survey was succeeded by the National Survey for Wales, but this has not so far featured a property survey. A new property survey is currently being planned and fieldwork is expected to begin in 2017.

6.5 The BEIS annual report on Fuel Poverty Statistics 2016 includes a brief section on fuel poverty in devolved administrations. Users are advised to contact BEIS for further information on fuel poverty and other energy policy issues across UK.²⁰

Geographies below England level

6.6 Frequent queries are received on the availability of EHS data for geographies below England level, such as at former Government Office Region (GOR) level, local authority level and lower super output layer (LSOA). Unfortunately, due to the relatively small sample size at local authority and LSOA level (the EHS is not designed to be representative at local authority level), DCLG has concerns about both the quality of estimates that could be achieved and the potential disclosure of individual respondents if these datasets were published with no restrictions. Restricted access to these variables is available by Secure Access licence at UK Data Service (see Accessibility section).

6.7 The EHS datasets available on End User Licence from UK Data Service provide geographical identifier for statistical region (England divided into nine areas) to allow users to create their own regional statistics. As outlined in the Written Ministerial Statement of 18 September 2012, *Official Report*, Column 32WS, DCLG does not publish official statistics by government office region.

Harmonised questions

6.8 A cross-governmental programme of work is currently underway looking into standardising inputs and outputs for use in National Statistics. This is known as harmonisation. The aim is to make it easier for users to draw clearer and more robust comparisons between data sources. The EHS adopts harmonised questions where possible.²¹

¹⁸ http://www.nihe.gov.uk/index/corporate/housing_research/house_condition_survey.htm

¹⁹ <http://wales.gov.uk/statistics-and-research/living-in-wales-survey/?lang=en>

²⁰ <https://www.gov.uk/government/statistics/annual-fuel-poverty-statistics-report-2016>

²¹ <http://www.ons.gov.uk/ons/guide-method/harmonisation/harmonisation-index-page/index.html>

Sources for further information or advice

- Link to headline report
<https://www.gov.uk/government/statistics/english-housing-survey-2014-to-2015-headline-report>
- Link to annual reports
<https://www.gov.uk/government/statistics/english-housing-survey-2014-to-2015-adaptations-and-accessibility-of-homes-report>
<https://www.gov.uk/government/statistics/english-housing-survey-2014-energy-report>
<https://www.gov.uk/government/statistics/english-housing-survey-2014-to-2015-first-time-buyers-and-potential-home-owners-report>
<https://www.gov.uk/government/statistics/english-housing-survey-2014-housing-and-well-being-report>
<https://www.gov.uk/government/statistics/english-housing-survey-2014-to-2015-housing-for-older-people-report>
<https://www.gov.uk/government/statistics/english-housing-survey-2014-to-2015-private-rented-sector-report>
<https://www.gov.uk/government/statistics/english-housing-survey-2014-to-2015-housing-stock-report>
<https://www.gov.uk/government/statistics/english-housing-survey-2014-to-2015-smoke-alarms-in-english-homes-report>
<https://www.gov.uk/government/statistics/english-housing-survey-2014-to-2015-social-rented-sector-report>
- Link to technical report
<https://www.gov.uk/government/collections/english-housing-survey-technical-advice>
- Link to data security strategy
<https://www.gov.uk/government/publications/english-housing-survey-data-security-strategy-and-arrangements>
- Link to data sharing agreement
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/539739/DataSharing_Agreement_DCLG_DECC_May_2016_with_signatures.pdf
- Link to explanation of open data standards
<https://www.gov.uk/government/publications/open-data-white-paper-unleashing-the-potential>
- Guidelines for Measuring Statistical Quality of official statistics, published by the ONS, available at
<http://www.ons.gov.uk/ons/guide-method/method-quality/quality/guidelines-for-measuring-statistical-quality/index.html>
- Official Statistics Release Calendar
<https://www.gov.uk/government/statistics/announcements>
- Code of Practice for Official Statistics
<https://www.statisticsauthority.gov.uk/monitoring-and-assessment/code-of-practice/>
- Further information email: ehs@communities.gsi.gov.uk

The United Kingdom Statistics Authority has designated these statistics as National Statistics, in accordance with the Statistics and Registration Service Act 2007 and Signifying compliance with the Code of Practice for Official Statistics.

Designation can be broadly interpreted to mean that the statistics:

- meet identified user needs;
- are well explained and readily accessible;
- are produced according to sound methods, and
- are managed impartially and objectively in the public interest.

Once statistics have been designated as National Statistics it is a statutory requirement that the Code of Practice shall continue to be observed.

Crown copyright, 2016

Copyright in the typographical arrangement rests with the Crown.

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. To view this licence, <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/> or write to the Information Policy Team, The National Archives, Kew, London TW9 4DU, or email: psi@nationalarchives.gsi.gov.uk.

This document/publication is also available on our website at www.gov.uk/dclg

If you have any enquiries regarding this document/publication, complete the form at <http://forms.communities.gov.uk/> or write to us at:

Department for Communities and Local Government
Fry Building
2 Marsham Street
London
SW1P 4DF
Telephone: 030 3444 0000
Email: ehs@communities.gsi.gov.uk

For all our latest news and updates follow us on Twitter: <https://twitter.com/CommunitiesUK>

July 2016

ISBN: 978-1-4098-4878-3